Discovery of Activity Patterns using Topic Models

Paper by Tâm Huỳnh, Mario Fritz and Bernt Schiele

Presentation by Roland Meyer

Introduction

- Detect routines based on body movement
- Complex due to large variations in activities

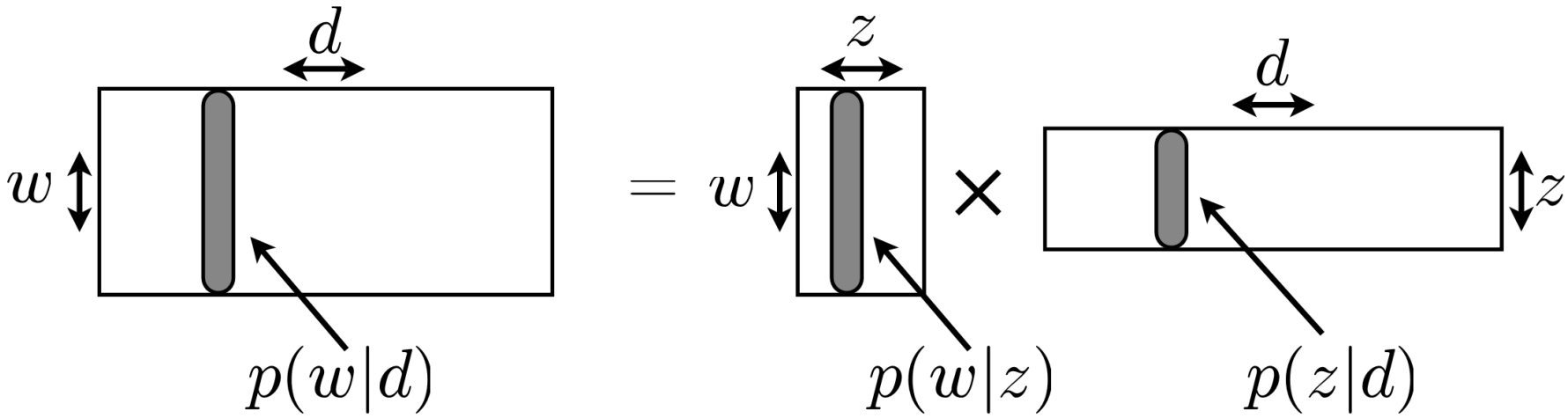
Contributions

- New method to recognize daily routines
- Reusing an established method from text processing
- Applicable without user annotation

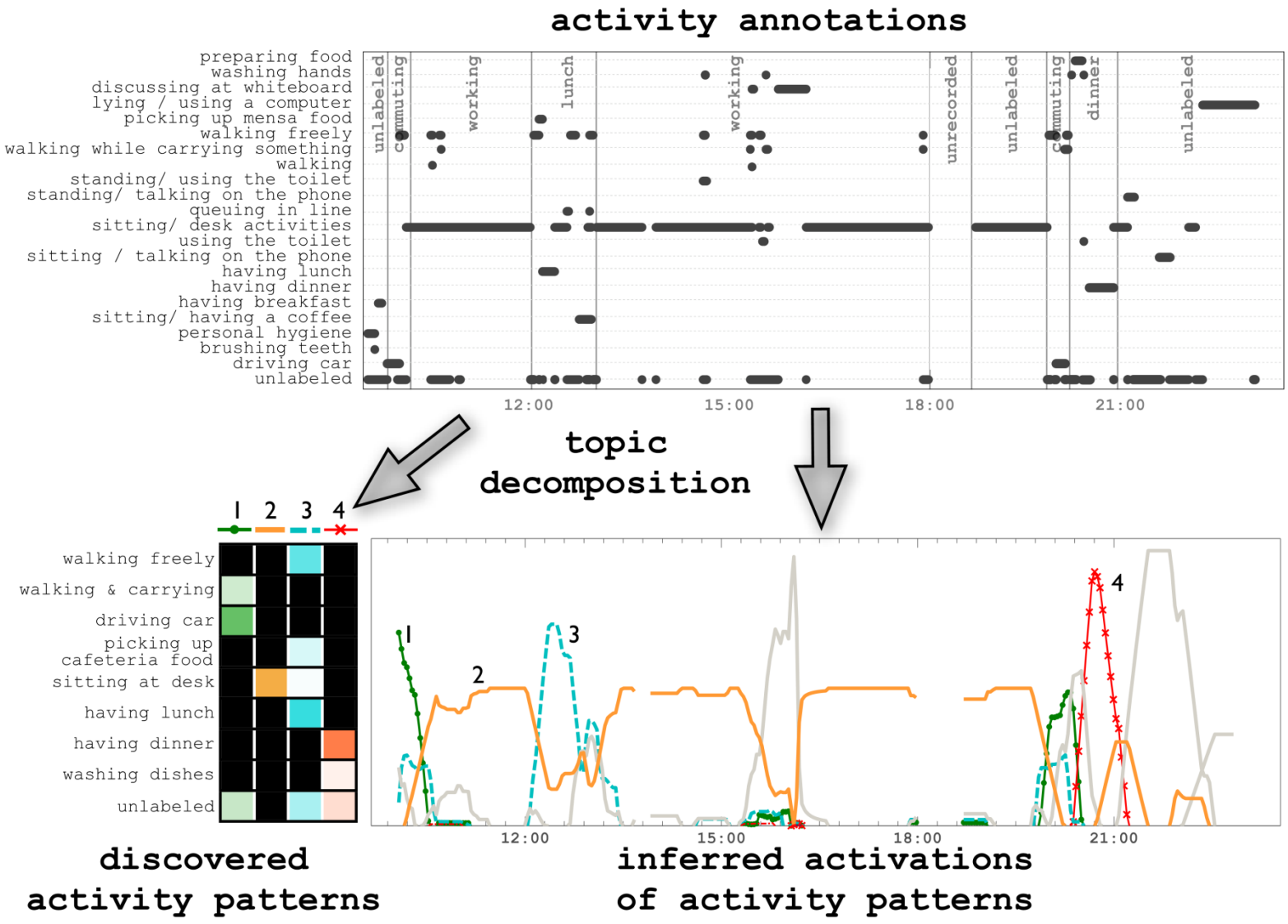
Topic Models

- Used for text processing for classification
- Collection of words (“Bag-of-words”)
- Unsupervised

Topic Models



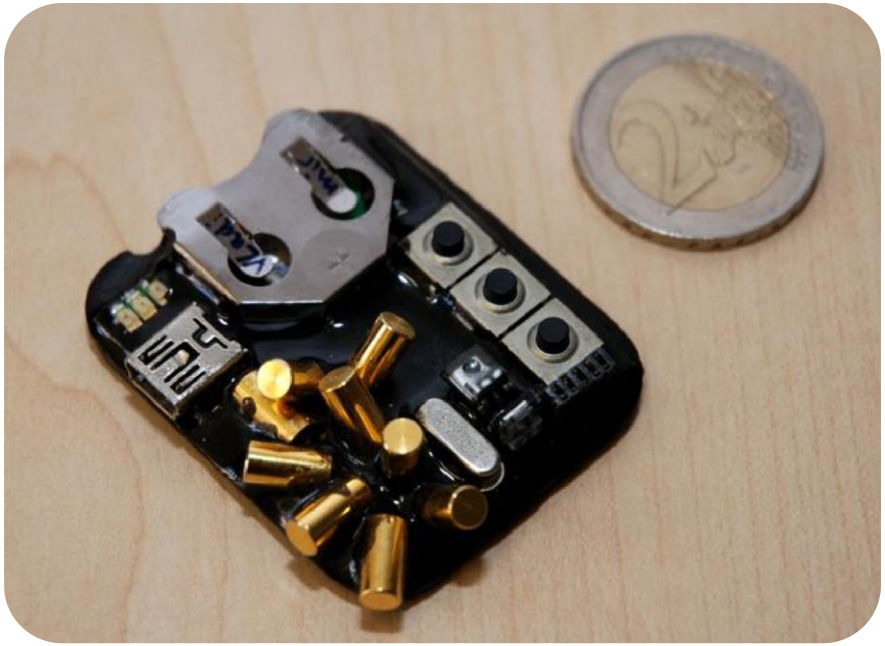
Daily Routine Modeling



Data collection

- 1 person
- 16 days
- 2 wearable sensors

- Accelerometer
- Realtime clock
- 4 hours of memory



Annotation

- Online annotation
  - Periodic set of questions on cell phone
  - Time diary
  - Occasional snapshots
- Offline annotation
  - User could correct / complement data
- Used as ground truth

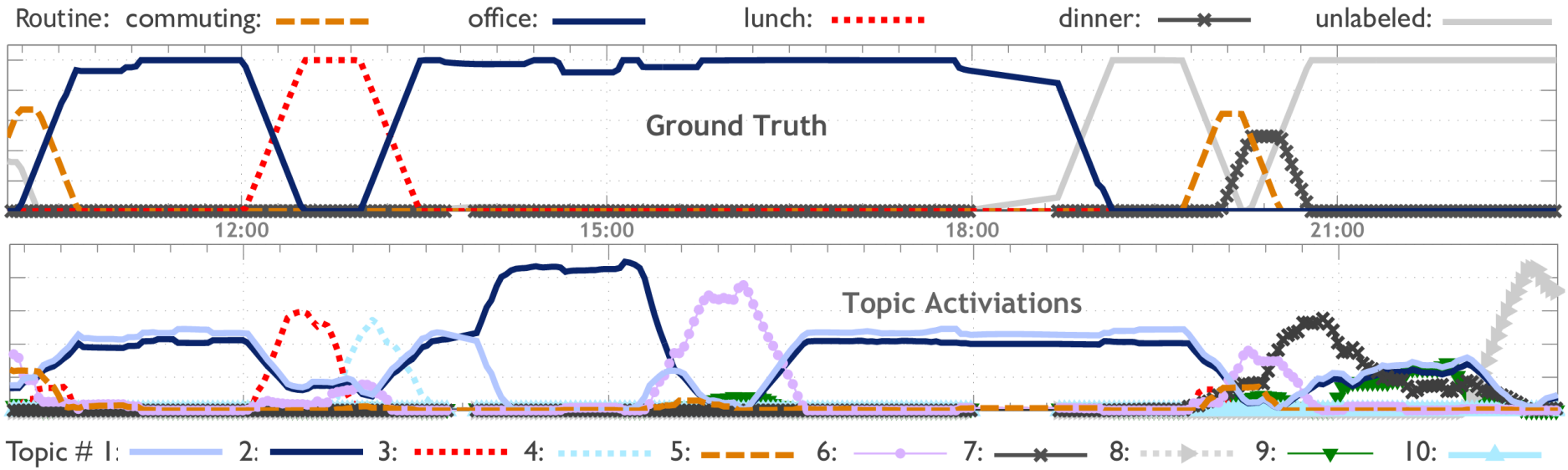
Discovering activities

- 34 distinct activities
- Mean, variance, frequency from acceleration sensors
- Combined with time-of-day
- SVMs, HMMs, Naive Bayes evaluated as classifiers
- 72.7% accuracy
- Great variations
- Problems with short and similar tasks

Discovering topics

- Latent Dirichlet Allocation on activity data
- Sliding window of 30 min. over activity stream
- 10 topics

Discovering topics



(a)

1 *sitting / desk activities (1.0)*
2 *sitting / desk activities (0.99)*
3 *having lunch (0.5), walking freely (0.26), picking up cafeteria food (0.06), queuing in line (0.06), unlabeled (0.04), walking & carrying sth (0.02), brushing teeth (0.02)*
4 *standing / using toilet (0.69), walking freely (0.18), queuing in line (0.06), walking (0.02), desk activities (0.02)*
5 *driving car (0.33), walking & carrying sth. (0.21), sitting / desk activities (0.2), walking (0.14), unlabeled (0.08)*
6 *using the toilet (0.69), walking freely (0.17), discussing at whiteboard (0.06), sitting / desk activities (0.03), brushing teeth (0.02)*
7 *having dinner (0.77), desk activities (0.1), washing dishes (0.08), unlabeled (0.04)*
8 *lying / using computer (1.0)*
9 *unlabeled (0.87), driving bike (0.04), washing dishes (0.02), stand/ use toilet (0.02), washing hands (0.02), standing / using the phone (0.02,*
10 *watching a movie (1.0)*

(b)

Results on Discovering topics

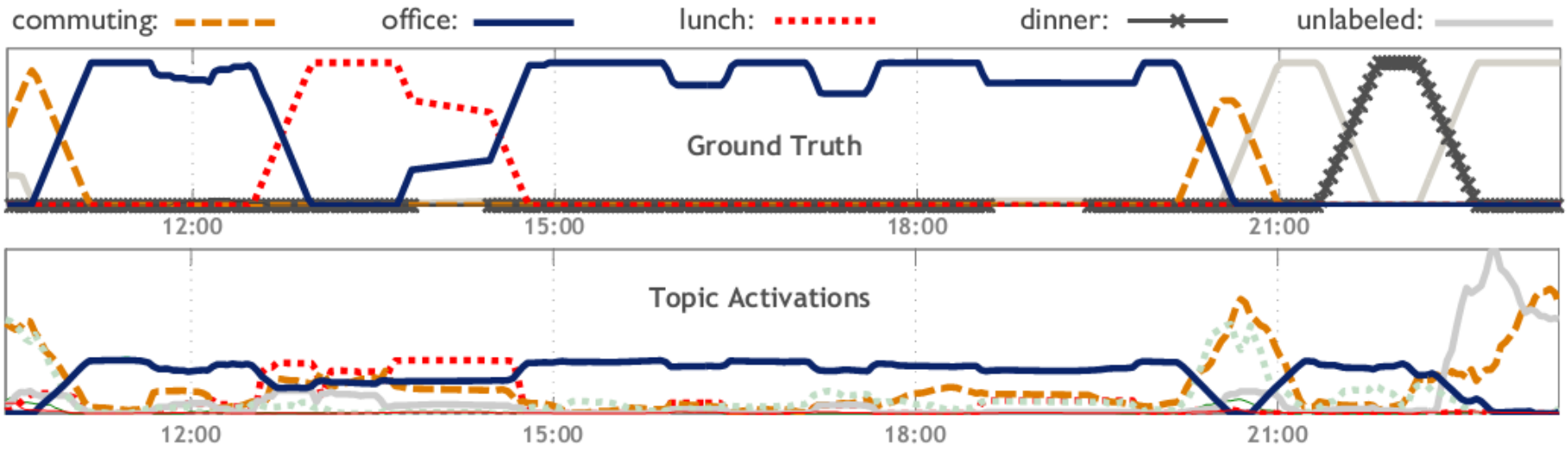
- Precision and recall calculated for 6 of 7 day to cross-validate results
- Supervised classifier using HMMs to calculate baseline

Routine	*Correlation*	*Precision*	*Recall*
Dinner	0.7	75.5	40.2
Commuting	0.6	85.5	51.8
Lunch	0.8	87.0	83.3
Office Work	0.8	96.4	93.7
Mean	0.7	86.1	67.2

Unsupervised learning

- Get rid of user annotations
- Labels from data clustering

Routine	*Correlation*	*Precision*	*Recall*
Dinner	0.6	56.9	40.2
Commuting	0.5	83.5	71.1
Lunch	0.8	73.8	70.2
Office Work	0.6	93.4	81.8
Mean	0.6	76.9	65.8



Future work

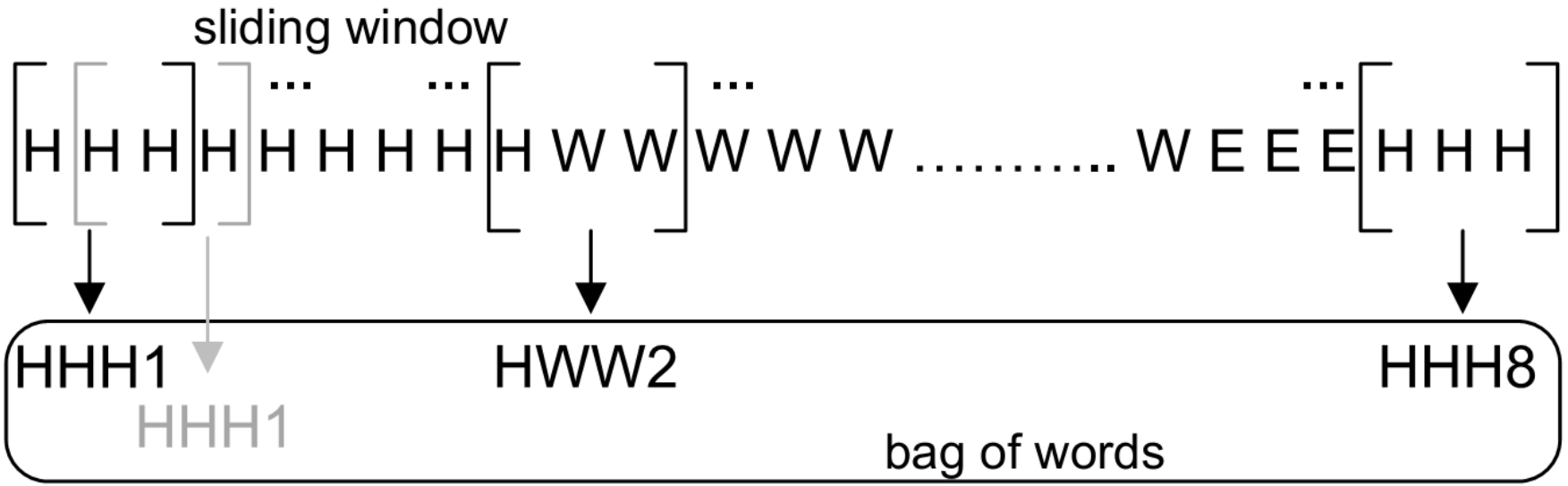
- Semi-supervision
- Noise modeling
- Include location information

- More users with more diverse lives
- Build applications
- Use better sensors (more memory)

Including location

- “Discovering Daily Routines from Google Latitude with Topic Models” by Laura Ferrari and Marco Mamei
- “Discovering Human Routines from Cell Phone Data with Topic Models” by Katayoun Farrahi and Daniel Gatica-Perez

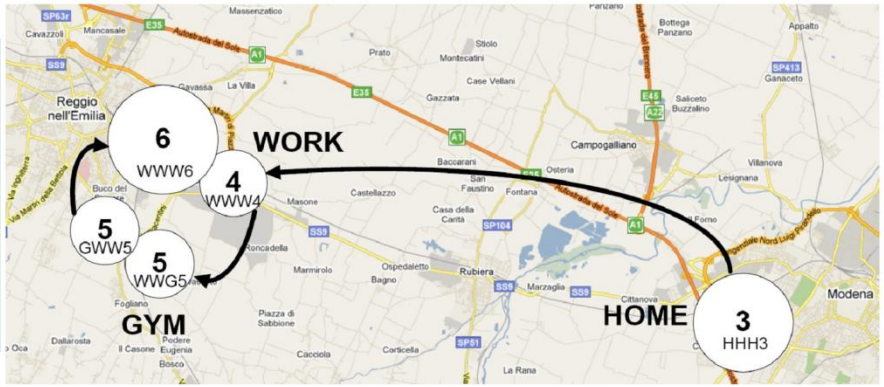
Including location

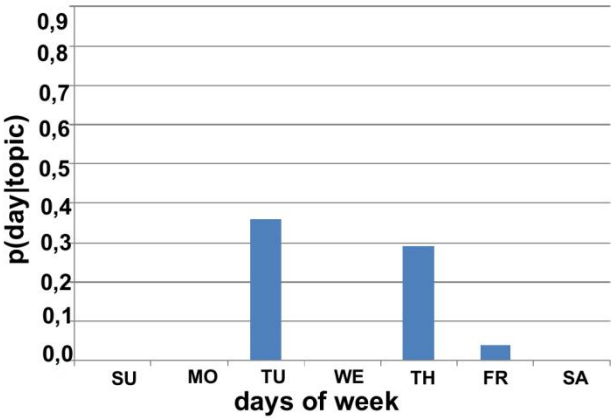


"Discovering Daily Routines from Google Latitude with Topic Models" - Laura Ferrari and Marco Mamei

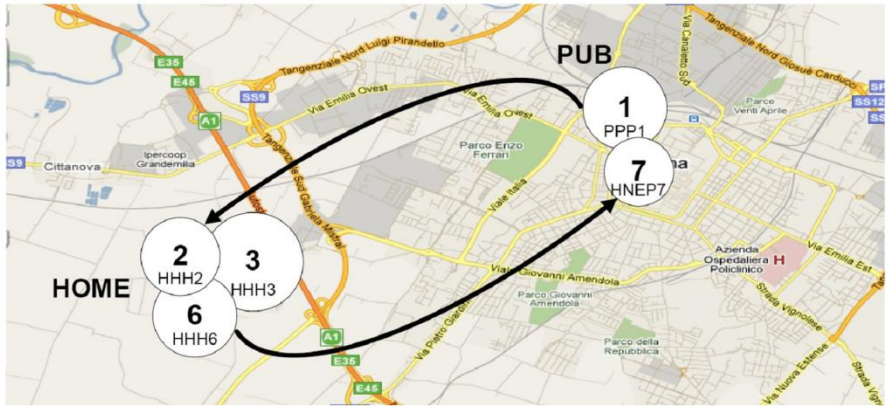
Including location

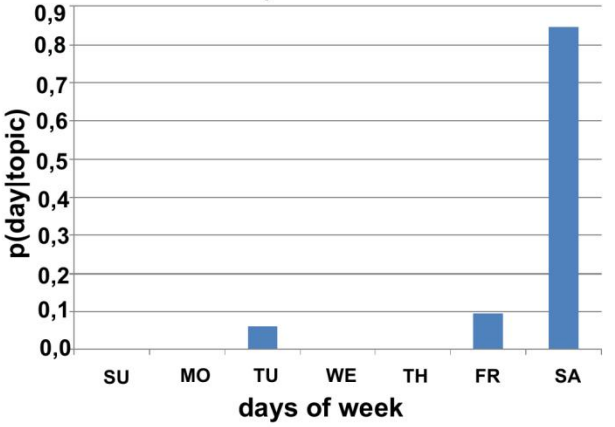
TOPIC 0	
word	p(w|z)
W W W 6	0,16
H H H 3	0,12
W W W 4	0,08
W W G 5	0,03
G W W 5	0,03



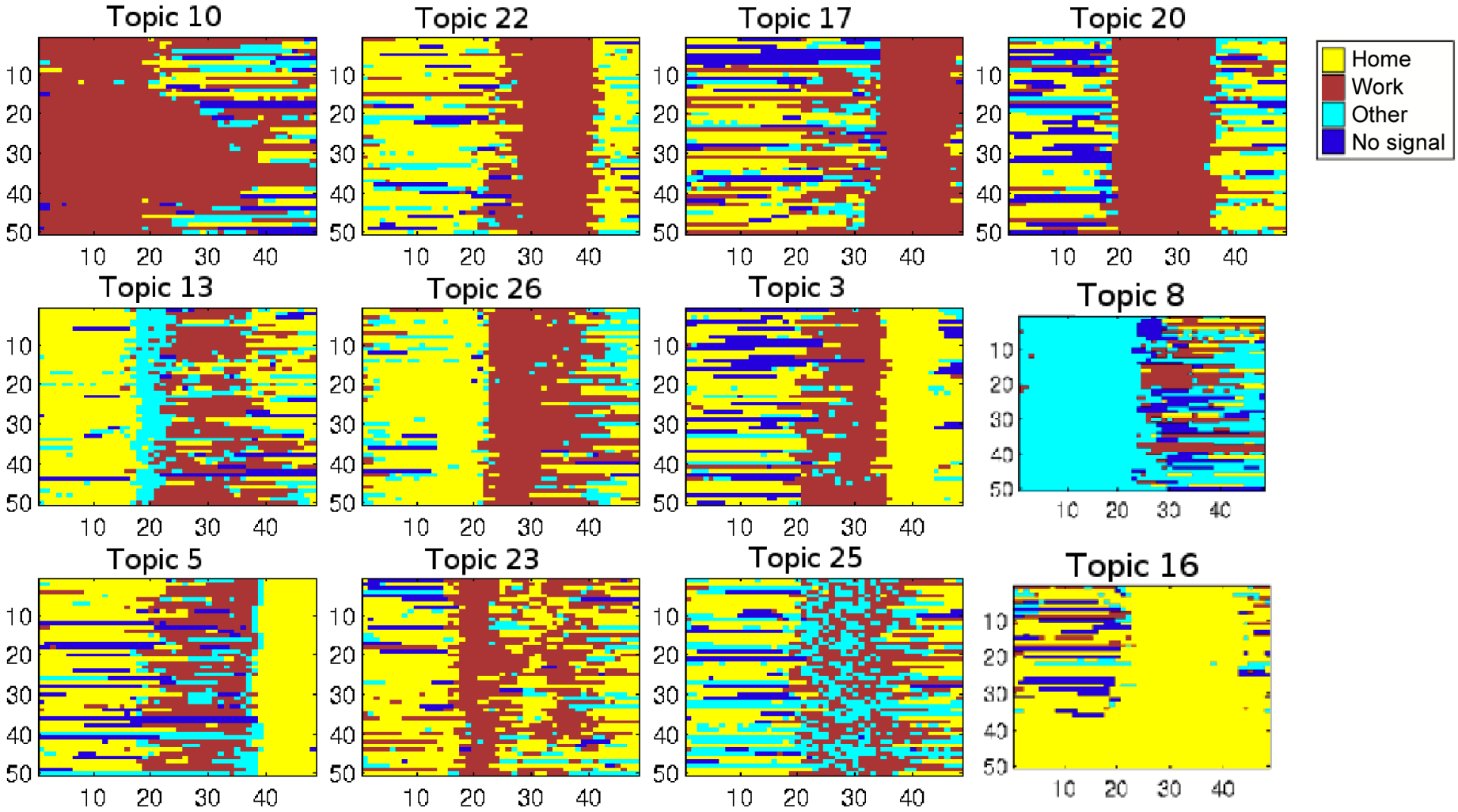


TOPIC 21	
word	p(w|z)
H H H 3	0,12
P P P 1	0,10
H H H 6	0,10
H H H 2	0,07
H NE P 7	0,02





Including location



"Discovering Human Routines from Cell Phone Data with Topic Models" - Katayoun Farrahi and Daniel Gatica-Perez

Reviews

- Average score: 1.75 (accept)

- Solid ground truth
- Privacy not addressed
- Spelling errors, graphs badly placed
- No automation, data needs to be manually copied